Faithful mRNA splicing depends on the Prp19 complex subunit faint sausage and is required for tracheal branching morphogenesis in Drosophila

Sauerwald, Julia ; Soneson, Charlotte ; Robinson, Mark D ; Luschnig, Stefan

Abstract: Morphogenesis requires the dynamic regulation of gene expression, including transcription, mRNA maturation and translation. Dysfunction of the general mRNA splicing machinery can cause surprisingly specific cellular phenotypes, but the basis for these effects is not clear. Here we show that the Drosophila faint sausage (fas) locus, implicated in epithelial morphogenesis and previously reported to encode a secreted immunoglobulin domain protein, in fact encodes a subunit of the spliceosome-activating Prp19 complex, which is essential for efficient pre-mRNA splicing. Loss of zygotic fas function globally impairs the efficiency of splicing, and is associated with widespread retention of introns in mRNAs and dramatic changes in gene expression. Surprisingly, despite these general effects, zygotic fas mutants show specific defects in tracheal cell migration during mid-embryogenesis when maternally supplied splicing factors have declined. We propose that tracheal branching, which relies on dynamic changes in gene expression, is particularly sensitive for efficient spliceosome function. Our results reveal an entry point to study requirements of the splicing machinery during organogenesis and provide a better understanding of disease phenotypes associated with mutations in general splicing factors.

DOI: https://doi.org/10.1242/dev.144535

Development Advance Online Articles. First posted online on 13 January 2017 as 10.1242/dev.144535
Access the most recent version at http://dev.biologists.org/lookup/doi/10.1242/dev.144535

Faithful mRNA splicing depends on the Prp19 complex subunit *faint sausage* and is required for tracheal branching morphogenesis in *Drosophila*

Julia Sauerwald[1, 2], Charlotte Soneson[2,3], Mark D. Robinson[2,3] and Stefan Luschnig[1, 2]

[1] Institute of Neurobiology, University of Münster, Badestrasse 9, 48149 Münster, Germany; Cluster of Excellence EXC 1003, Cells in Motion (CiM), 48149 Münster, Germany

[2] Institute of Molecular Life Sciences, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

[3] SIB Swiss Institute of Bioinformatics, Zürich, Switzerland

Correspondence: luschnig@uni-muenster.de

Key words: *Drosophila melanogaster*, tracheal system, branching morphogenesis, mRNA splicing, spliceosome, *faint sausage*, *fandango*

ABSTRACT

Morphogenesis requires the dynamic regulation of gene expression, including transcription, mRNA maturation and translation. Dysfunction of the general mRNA splicing machinery can cause surprisingly specific cellular phenotypes, but the basis for these effects is not clear. Here we show that the *Drosophila faint sausage* (*fas*) locus, implicated in epithelial morphogenesis and previously reported to encode a secreted immunoglobulin domain protein, in fact encodes a subunit of the spliceosome-activating Prp19 complex, which is essential for efficient pre-mRNA splicing. Loss of zygotic *fas* function globally impairs the efficiency of splicing, and is associated with widespread retention of introns in mRNAs and dramatic changes in gene expression. Surprisingly, despite these general effects, zygotic *fas* mutants show specific defects in tracheal cell migration during mid-embryogenesis when maternally supplied splicing factors have declined. We propose that tracheal branching, which relies on dynamic changes in gene expression, is particularly sensitive for efficient spliceosome function. Our results reveal an entry point to study requirements of the splicing machinery during organogenesis and provide a better understanding of disease phenotypes associated with mutations in general splicing factors.

INTRODUCTION

mRNA splicing is required to process nearly all eukaryotic transcripts. As splicing can be rate-limiting for efficient gene expression (Guilgur *et al*., 2014; Hoyle and Ish-Horowicz, 2013), reduced splicing efficiency can perturb many cellular processes. Although defects in RNA processing have been associated with pleiotropic phenotypes during development (Golling *et al*., 2002), inactivation of splicing factors can cause surprisingly specific cellular defects (Chen *et al*., 1998; van der Lelij *et al*., 2014). Mutations in core components of the spliceosome are associated with human diseases including Retinitis pigmentosa and Spinal muscular atrophy (Faustino, 2003), and can contribute to carcinogenesis (Quesada *et al*., 2011). However, the basis for these specific phenotypes is not clear.

Morphogenesis requires dynamically regulated gene expression programs to coordinate cell movements and differentiation. Tracheal branching morphogenesis in *Drosophila* is guided by the dynamic expression of the Fibroblast Growth Factor (FGF) Branchless (Bnl; Sutherland *et al*., 1996) in cells surrounding the tracheal primordia. Bnl activates the FGF receptor (FGFR) Breathless (Btl) on tracheal cells, which move as a cohort towards the Bnl source. Here we show that mutations in the *faint sausage* (*fas*) gene (Nüsslein-Volhard *et al*., 1984) specifically affect branch outgrowth and cellular rearrangements during tracheal morphogenesis. We found that *fas* encodes a subunit of the spliceosome-activating Prp19 complex (Prp19C; also known as NineTeen complex, NTC; Chanarat and Sträßer, 2013), contrary to an earlier report that *fas* encodes a secreted immunoglobulin (Ig) domain protein (Lekven *et al*., 1998). Lack of zygotic *fas*/*Prp19C* function broadly impairs the efficiency of mRNA splicing and leads to extensive changes in global gene expression. Our findings suggest that tracheal branching morphogenesis is particularly sensitive for efficient spliceosome function, thus providing an entry point to investigate the requirements of splicing during organogenesis.

MATERIALS AND METHODS

Drosophila strains and genetics

Unless noted otherwise, *Drosophila* stocks are described in FlyBase and were obtained from the Bloomington Stock center: btl^{724} (Ghabrial and Krasnow, 2006), fas^{1} (Nüsslein-Volhard *et al*., 1984), $CG6197^{+}$ (fas^{+}, genomic *fas* transgene), *UAS-CG6197-myc (UAS-Fas-myc)*, $fand^{1}$, $fand^{2}$ (Guilgur *et al*., 2014), *UAS-bnl* (Sutherland *et al*., 1996), *UAS-λBtl* (Lee *et al*., 1996), *UAS-palm-mNeonGreen* (this study), *UAS-CG17716* (this study), *btl-Gal4* (Shiga *et al*., 1996), *Abd-B-Gal4* (de Navas *et al*., 2006), *da-Gal4, 69B-Gal4, btl-Gal80* (Nikolova and Metzstein, 2015). fas^{H124} and fas^{P218} were isolated in an ethyl methanesulfonate (EMS) mutagenesis screen (Förster *et al*., 2010; Supplementary Material).

Immunoblots

Protein extracts from embryos (15-18 hours after egg lay; hAEL) were analyzed on immunoblots. Antibodies were rabbit anti-Fas (1:1,000; Guilgur *et al*., 2014), mouse anti-alpha-Tubulin Dm1A (1:10,000; Sigma), goat anti-rabbit Superclonal HRP conjugate (1:5,000; Thermo Fischer), and goat anti-mouse Superclonal HRP conjugate (1:5,000; Thermo Fischer). Three independent replicates (embryo collections) were analyzed.

Molecular Biology

A DNA fragment corresponding to *CG17716-PA* cDNA (FlyBase) was synthesized (GenScript), cloned into pUAST-attB (*Eco*RI/*Hind*III), and integrated into the attP2 (68A4) landing site using PhiC31 integrase (Bischof *et al*., 2007).

The coding sequence of monomeric yellow-green fluorescent protein (mNeonGreen; Shaner *et al*., 2013) fused to an N-terminal palmitoylation signal was synthesized (GenScript) using the *D. melanogaster* codon distribution, cloned into pUASt-attB (*Eco*RI/*Xba*I), and integrated into the attP2 landing site.

Immunostainings

Embryos were fixed in 4% formaldehyde for 20 min and devitellinized in methanol/heptane. Primary antibodies were chicken anti-GFP (1:500; Abcam #13970), mouse anti-DSRF (1:300; Samakovlis *et al*., 1996), rabbit anti-Sal (1:40; Kuhnlein and Schuh, 1996), mouse anti-Tango

(1:100; DSHB), mouse anti-myc (9E10; 1:300; DSHB) and rabbit anti-dpERK (1:100; Cell Signaling Technology #4370). Goat secondary antibodies were conjugated with DyLight 488 (1:500; Abcam), Alexa Fluor 568 (1:300; Molecular Probes) or Cy5 (1:500; Jackson ImmunoResearch). Chitin was detected as previously described (Caviglia and Luschnig, 2013). Neurons were labeled using anti-HRP (1:1,000; Dianova) conjugated with Alexa Fluor 647.

Imaging

Imaging was performed on an Olympus FV1000 confocal microscope with 20x/0.75 NA, 40x/1.3 NA and 60x/1.35 NA objectives or on a Zeiss LSM710 with a 20x/1.0 NA objective. For live imaging, staged embryos were dechorionated, glued on a coverslip and immersed in Voltalef 10S oil. Images were processed using ImageJ (v2.0.0), Imaris (v8.3.1; Bitplane) and Adobe Photoshop. At least three embryos per genotype were analyzed.

RNA Sequencing

Total RNA was isolated from 12-13 hAEL *fas*^{P218} *btl-Gal4 UAS-GFP UAS-Verm-mRFP* embryos and from control embryos (carrying the parental chromosome) using TRIzol (Thermo Fischer). RNA was precipitated in isopropanol with 0.3 M sodium acetate and treated with DNase I (Ambion) for 25 min at 37°C. cDNA libraries were generated using the Illumina RNASeq protocol and were sequenced with an Illumina HiSeq instrument. RNASeq experiments in three biological replicates (independent embryo collections) of control and *fas*^{P218} embryos yielded 27 to 69 million paired-end reads per sample. Details on RNA-Seq data analyses are provided in Supplementary Materials and Methods.

RESULTS

A new gene required for tracheal branching morphogenesis

We isolated two allelic embryonic lethal mutants (*P218* and *H124*) that were defective in primary tracheal branching (Fig. 1). While wild-type embryos have formed an interconnected tracheal network by 12 hAEL, tracheal branch outgrowth was impaired in *P218* and *H124* homozygous (Fig. 1A,B; Supplementary Movie 1) and in *P218*/*H124* trans-heterozygous embryos. Primary branching was severely reduced, but not entirely abolished as in *btl* mutants, which lack the FGF receptor (Fig. 1C-E). *P218* and *H124* tracheal primordia remained elongated sacs with dorsal extensions that frequently detached from the remaining primordia (Fig. 1B; Supplementary Movie 2). Adjacent metameres often formed a partially interconnected dorsal trunk (DT; Fig. 1B) that appeared to become stretched during tube elongation, especially at sites where relatively few cells were present (Fig. 1B; Supplementary Movie 3). Persisting DT connections revealed deposition of luminal material (Fig. 1F,G). While tracheal development was severely affected, earlier processes, including gastrulation and germband retraction (Fig. 1A,B), appeared normal. However, the mutant embryos showed defects during later embryogenesis, including abnormal dorsal closure and head involution (not shown; Liu *et al*., 1999).

H124 and *P218* are allelic to the *faint sausage* (*fas*) locus

We mapped the lethality of *H124* and *P218* to a small interval (50B4-B6) comprising six annotated genes (Fig. 2A; Supplementary Materials and Methods). A lethal P-element insertion *(P{PZ}05488)* in this interval failed to complement *H124* and *P218* mutants. *P{PZ}05488* is an allele of the *faint sausage* (*fas*; Nüsslein-Volhard *et al*., 1984) locus, which was reported to encode a secreted immunoglobulin domain protein (CG17761) expressed in the central nervous system (CNS) and in clusters of epidermal cells (Lekven *et al*., 1998). Furthermore, the EMS-induced fas^{1} mutation (Nüsslein-Volhard *et al*., 1984) failed to complement *P218* and *H124*, indicating that the two mutations are allelic to *fas*. However, expression of a UAS-*CG17761* transgene corresponding to the reported *fas* locus (Lekven *et al*., 1998) in the ectoderm of *fas* embryos did not lead to noticeable rescue of the tracheal defects (data not shown).

fas mutations affect CG6197, a subunit of the spliceosomal Prp19 complex, and not the immunoglobulin domain protein CG17716

P{PZ}05488 is inserted upstream of the *fandango (fand, CG6197*) gene, which encodes a subunit of the NineTeen Complex/Prp19 complex (NTC/Prp19C; Fig. 2A; Guilgur *et al*., 2014) that is essential for spliceosome activation (Chanarat and Sträßer, 2013). Surprisingly, two *fand* alleles, $fand^{1}$ and $fand^{2}$ (Guilgur *et al*., 2014), failed to complement *P218* and *H124* mutants. Consistent with these findings, sequence analysis revealed premature stop codons in the *CG6197*/*fand* gene in *H124* and *P218* mutants (Fig. 2A). Moreover, we found that the fas^{1} allele (Nüsslein-Volhard *et al*., 1984) carries a premature stop codon in *CG6197* within the same codon as *P218* (Fig. 2A). Conversely, we did not find any missense or nonsense mutations in the *CG17716* coding sequence in *H124*, *P218* and fas^{1} mutants. Finally, the tracheal and CNS defects of fas^{H124}/fas^{1} embryos were completely rescued by a genomic construct containing a wild-type copy of *CG6197* (Fig. 2C-F; Guilgur *et al*., 2014). Together, these findings indicate that the *faint sausage* phenotype is caused by mutations in the NTC/Prp19C subunit CG6197 also known as Fand (Guilgur *et al*., 2014), and not, as previously reported, in the immunoglobulin domain protein CG17716 (Lekven *et al*., 1998). We therefore refer to the *CG6197* gene as *faint sausage* (*fas*) from here on.

Fas is required outside of tracheal cells for tracheal cell migration

To investigate whether *fas* function is required in tracheal cells for their normal migration we expressed a myc-tagged Fas (Fas-myc) construct in tracheal cells of *fas* mutants. However, tracheal-specific expression of Fas-myc was not sufficient to rescue tracheal morphogenesis (Fig. 2G,G'). Conversely, tracheal development was rescued when Fas-myc was expressed in all cells within an entire body segment (Fig. 2H,H'), Furthermore, expression of Fas-myc throughout the embryo except for tracheal cells (using ubiquitous *da*-Gal combined with *btl*-Gal80 to inhibit Gal4 specifically in tracheal cells) led to nearly complete rescue of tracheal development in *fas* embryos (Fig. 2I-J'). Although we cannot exclude that *btl*-Gal80-mediated inhibition of *da*-Gal4-driven Fas-myc expression was incomplete, these findings suggest that tracheal cell migration depends on *fas* function in surrounding tissues.

Zygotic loss of *fas* causes widespread intron retention

Maternally provided Fas protein is essential for efficient mRNA splicing in the early embryo (Guilgur *et al*., 2014; Martinho *et al*., 2015), but zygotic functions of *fas* have not been addressed thus far. We found that Fas protein levels were strongly reduced in fas^{H124} and fas^{P218} late-stage embryos (15-18 hAEL), although residual amounts of Fas protein, presumably representing maternal Fas protein, were still detectable by immunoblot (Fig. 2B).

To systematically identify changes in mRNAs, which may account for the tracheal defects in *fas* embryos, we performed high-throughput transcriptome sequencing (RNAseq). Total RNA was extracted from fas^{P218} and control embryos at 12-13 hAEL, when tracheal development showed first signs of perturbation in *fas* mutants. Most strikingly, we observed an accumulation of transcripts with retained intronic sequences in *fas* mutants (Fig. 3A). Out of 47,590 introns present in 11,059 genes, the relative inclusion rate of 6,453 introns (13.6%) was increased significantly (FDR < 0.01) and at least 2.83-fold in fas^{P218} embryos (Fig. 3B; log2 FC ≥ 1.5; Supplementary Table 1). These retained introns were present in 3,629 genes, suggesting that lack of Fas does not uniformly affect splicing of all introns, but that a subset of transcripts or introns is more susceptible to escape splicing in the absence of Fas. Similarly, among the set of retained introns, some were more frequently retained than others. This was observed also for different introns within the same transcript (*e.g*. CG17716; Fig. 3B,C). For example, of the 6,453 introns with at least 2.83-fold increased retention (FDR < 0.01, log2 FC ≥ 1.5), 2,968 were increased fourfold or more (log2 FC ≥ 2) and 896 were increased at least sixteen-fold (log2 FC ≥ 4), suggesting that certain introns may be especially prone to splicing defects when Prp19C function becomes limiting. While introns with high retention rates show some association with intron length, GC-content, number of introns per gene and transcript expression level, these weak associations could be partially explained by higher statistical power when counts are higher (Supplemental Figure 1).

Out of 326 genes with annotated functions during tracheal development (GO:0007424; FlyBase), 214 genes were considered as expressed (Supplementary Materials and Methods). 90% (193 genes) of these expressed tracheal genes revealed abnormal transcript processing (gene-level FDR<0.01; Fig. 3D). Most of the remaining normally processed transcripts were derived either from intron-less genes or from genes with only few short introns (Fig. 3D). The tracheal cell migration defects in *fas* embryos could in principle be explained by abnormal splicing of transcripts encoding components of FGF and EGF signaling (Fig. 3D). However, transcripts with highest intronic fold-changes in *fas* mutants include several receptor tyrosine kinase (RTK) signaling components acting either upstream (Bnl, Sfl) or downstream (Stumps,

Sos) of the RTKs (Fig. 3E), suggesting that the tracheal defects in *fas* embryos are unlikely to be attributable to changes in the expression of a specific gene.

In addition to the splicing defects, many genes showed dramatic changes in transcript levels (Supplementary Table 2). Of 9,606 expressed genes 11.5% were downregulated at least 0.35-fold (log2 FC $\leq$ -1.5; FDR < 0.01). 6.7% were upregulated at least 2.83-fold (log2 FC $\geq$ 1.5; FDR < 0.01). We conclude that lack of zygotic Fas protein severely affects the efficiency of RNA splicing in the developing embryo, resulting in qualitative (intron retention) as well as quantitative (transcript abundance) changes in gene expression.

EGF and FGF signaling are compromised in *fas* embryos

Analysis of all mis-spliced tracheal transcripts in *fas* mutants revealed abnormal transcript processing and changes in expression levels of several components of EGF and FGF signaling (Supplementary Figure 2A,B). To validate the functional consequences of the splicing defects, we analyzed EGF and FGF signaling activity during tracheal morphogenesis. The branch identity gene *spalt* (*sal;* Kuhnlein and Schuh, 1996) is expressed in tracheal placodes upon activation of Wnt and EGF signaling (Chihara and Hayashi, 2000). Sal signals in tracheal cells of *fas* embryos were strongly reduced (Fig. 4A,B). Similarly, FGF-dependent ERK activation (dpERK) at tracheal branch tips was severely diminished (Fig. 4C,D), and FGF-induced specification of DSRF-positive tracheal terminal cells was completely abolished (Fig. 4E,F). However, the residual tracheal branching in *fas* mutants (Fig. 1E) suggests that FGF and EGF signaling pathways are partially active, but that their output is quantitatively reduced. We therefore asked whether the output of FGF signaling can be restored in *fas* mutants through expression of a constitutively active FGF receptor (λBtl; Lee *et al*., 1996). Indeed, we observed high ERK activation and DSRF-positive terminal cells upon constitutive activation of Btl FGFR signaling in tracheal cells of *fas* embryos (Fig. 4G-I).

The requirement of *fas* outside tracheal cells suggested that events upstream of FGFR, possibly including the production of Bnl FGF, may be compromised in the absence of *fas* function. We therefore tested whether misexpression of Bnl FGF was able to restore tracheal branching in *fas* mutants. Misexpression of a *bnl* cDNA in the epidermis caused ectopic tracheal branching in wild-type controls, but not in *fas* embryos (Fig. 4J,K). Together, these findings suggest that while zygotic Fas function is not strictly required for signaling downstream of FGFR, it is essential for the activation of signaling by Bnl FGF, although the

tracheal branching defects in *fas* mutants are not solely attributable to mis-splicing of *bnl* transcripts.

DISCUSSION

We described the requirement of the NTC/Prp19C subunit Fas for tracheal branching morphogenesis. Although mRNA splicing is generally required for transcript maturation, we found that embryos lacking zygotic *fas* function display surprisingly specific organogenesis defects. First, we show that the NTC/Prp19C subunit CG6197 is required for tracheal branching. We demonstrate that the *faint sausage* (*fas*) locus (Nüsslein-Volhard *et al*., 1984), previously reported to encode the secreted immunoglobulin domain protein CG17716 (Lekven *et al*., 1998), in fact encodes the NTC/Prp19C subunit CG6197 also known as *fandango* (*fand*; Guilgur *et al*., 2014). Second, we show that loss of zygotic *fas* function results in widespread perturbation of splicing, consistent with previous work demonstrating an essential requirement of maternal *CG6197/fand* function for efficient splicing during early embryogenesis (Guilgur *et al*., 2014; Martinho *et al*., 2015). Abnormal transcript processing in *fas* mutants manifests predominantly in intron retention, accompanied by changes in transcript abundance. Third, we show that compromised FGF and EGF signaling may contribute to the tracheal branching defects in *fas* mutants. The requirement of Fas in non-tracheal cells (Fig. 2G-J) and the results of our epistatic analysis (Fig. 4G-K) suggest that *fas* function is essential for the activation of Btl FGFR.

The late onset and the specific nature of tracheal and CNS (Lekven *et al*., 1998; Liu *et al*., 1999) defects in zygotic *fas* embryos was surprising, given that pre-mRNA processing is generally perturbed in the mutants. However, maternally provided gene products, including Fas protein itself, are likely to allow for largely normal development during early embryogenesis. As maternal Fas protein decays over time, Fas levels appear to become limiting at the onset of tracheal morphogenesis during mid-embryogenesis. Perturbed pre-mRNA processing in *fas* mutants is accompanied by substantial changes in the levels of many transcripts. Nonsense-mediated mRNA decay (Wilusz *et al*., 2001) is expected to degrade a large fraction of mis-spliced transcripts. In addition, indirect effects on transcriptional regulation are likely to influence gene expression in *fas* mutants. Of note, the abnormally processed transcripts in *fas* embryos include *CG17716* mRNA (Fig. 3C), consistent with the finding of Lekven *et al*. (1998) that CG17716 protein was undetectable in *fas* embryos.

We report an example of strikingly specific developmental defects associated with a lack of efficient splicing. How can a general perturbation of splicing lead to specific phenotypes? First, tissue-specific expression of some splicing factors, including Prp19 (Urano *et al*., 2006), might account for tissue-dependent differences in splicing efficiency. Second, pre-mRNAs contain diverse auxiliary *cis*-acting regulatory elements that are recognized by a multitude of splicing factors (Zhang *et al*., 2008). Consequently, different introns may show distinct sensitivities towards the lack of a given splicing factor. In addition, intron number, transcript abundance, and transcript stability may render some RNAs more prone to accumulating splicing errors than others. Finally, dynamic cellular processes, such as cellularization (Guilgur *et al*., 2014) and tracheal branching, which involve rapid regulation of gene expression, may depend more acutely on efficient mRNA processing. Consistent with this idea, highly expressed and rapidly regulated genes tend to have only few and short introns (Castillo-Davis *et al*., 2002; Jeffares *et al*., 2008).

Our findings allow to define the most sensitive gene expression events required for proper organogenesis. Thus, characterizing NTC/Prp19C function and regulation in different organs could contribute to a better understanding of how differential gene expression is regulated during organogenesis.

ACKNOWLEDGEMENTS

We thank Rui Gonçalo Martinho, the Developmental Studies Hybridoma Bank and the Bloomington *Drosophila* stock center for providing fly stocks and antibodies. The Functional Genomics Center Zurich (FGCZ) provided support with RNA sequencing experiments. We thank Dirk Beuchle for help with genetic mapping, Simone Mumbauer for generating *UAS-CG17716* flies, Wilko Backer for technical support, and Michaela Clever for comments on the manuscript. We are indebted to Christian Lehner for continuous support and discussions. We are grateful to Sofia Araujo and Rui Gonçalo Martinho for communication of unpublished work.

COMPETING INTERESTS

The authors declare no competing financial interests.

AUTHOR CONTRIBUTIONS

JS and SL conceived and designed the experiments. JS performed the experiments. JS, CS and MR analyzed RNA-seq data. JS and SL wrote the manuscript. All authors discussed the results and edited the manuscript.

FUNDING

JS was supported by a Boehringer Ingelheim Fonds fellowship. Work in SL's laboratory was supported by the Swiss National Science Foundation (SNF 31003A_141093_1), the University of Zurich, the Kanton Zurich, the "Cells-in-Motion" Cluster of Excellence (EXC 1003-CiM), and the University of Münster.

REFERENCES

Bischof, J., Maeda, R. K., Hediger, M., Karch, F. and Basler, K. (2007). An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 3312–3317.

Castillo-Davis, C. I., Mekhedov, S. L., Hartl, D. L., Koonin, E. V. and Kondrashov, F. A. (2002). Selection for short introns in highly expressed genes. *Nature Genetics* **31**, 415–418.

Caviglia, S. and Luschnig, S. (2013). The ETS domain transcriptional repressor Anterior open inhibits MAP kinase and Wingless signaling to couple tracheal cell fate with branch identity. *Development* **140**, 1240–1249.

Chanarat, S. and Sträßer, K. (2013). Splicing and beyond: the many faces of the Prp19 complex. *Biochim. Biophys. Acta* **1833**, 2126–2134.

Chen, E. J., Frand, A. R., Chitouras, E. and Kaiser, C. A. (1998). A link between secretion and pre-mRNA processing defects in Saccharomyces cerevisiae and the identification of a novel splicing gene, RSE1. *Mol. Cell. Biol.* **18**, 7139–7146.

Chihara, T. and Hayashi, S. (2000). Control of tracheal tubulogenesis by Wingless signaling. *Development* **127**, 4433–4442.

de Navas, L., Foronda, D., Suzanne, M. and Sánchez-Herrero, E. (2006). A simple and efficient method to identify replacements of P-lacZ by P-Gal4 lines allows obtaining Gal4 insertions in the bithorax complex of Drosophila. *Mech. Dev.* **123**, 860–867.

Faustino, N. A. and Cooper, T. A. (2003). Pre-mRNA splicing and human disease. *Genes & Development* **17**, 419–437.

Förster, D., Armbruster, K. and Luschnig, S. (2010). Sec24-dependent secretion drives cell-autonomous expansion of tracheal tubes in Drosophila. *Curr. Biol.* **20**, 62–68.

Ghabrial, A. S. and Krasnow, M. A. (2006). Social interactions among epithelial cells during tracheal branching morphogenesis. *Nature* **441**, 746–749.

Golling, G., Amsterdam, A., Sun, Z., Antonelli, M., Maldonado, E., Chen, W., Burgess, S., Haldi, M., Artzt, K., Farrington, S., et al. (2002). Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development. *Nature genetics* **31**, 135–140.

Guilgur, L. G., Prudêncio, P., Sobral, D., Liszekova, D., Rosa, A. and Martinho, R. G. (2014). Requirement for highly efficient pre-mRNA splicing during Drosophila early embryonic development. *Elife* **3**, e02181.

Hoyle, N. P. and Ish-Horowicz, D. (2013). Transcript processing and export kinetics are rate-limiting steps in expressing vertebrate segmentation clock genes. *Proc. Natl. Acad. Sci. U.S.A.* **110**, E4316–24.

Jeffares, D. C., Penkett, C. J. and Bähler, J. (2008). Rapidly regulated genes are intron poor. *Trends Genet.* **24**, 375–378.

Khodor, Y. L., Rodriguez, J., Abruzzi, K. C., Tang, C.-H. A., Marr, M. T. and Rosbash, M. (2011). Nascent-seq indicates widespread cotranscriptional pre-mRNA splicing in Drosophila. *Genes & Development* **25**, 2502–2512.

Kuhnlein, R. P. and Schuh, R. (1996). Dual function of the region-specific homeotic gene spalt during Drosophila tracheal system development. *Development* **122**, 2215–2223.

Lee, T., Hacohen, N., Krasnow, M. and Montell, D. J. (1996). Regulated Breathless receptor tyrosine kinase activity required to pattern cell migration and branching in the Drosophila tracheal system. *Genes & Development* **10**, 2912–2921.

Lekven, A. C., Tepass, U., Keshmeshian, M. and Hartenstein, V. (1998). faint sausage encodes a novel extracellular protein of the immunoglobulin superfamily required for cell migration and the establishment of normal axonal pathways in the Drosophila nervous system. *Development* **125**, 2747–2758.

Liu, X., Kiss, I. and Lengyel, J. A. (1999). Identification of genes controlling malpighian tubule and other epithelial morphogenesis in Drosophila melanogaster. *Genetics* **151**, 685–695.

Martinho, R. G., Guilgur, L. G. and Prudêncio, P. (2015). How gene expression in fast-proliferating cells keeps pace. *Bioessays* **37**, 514–524.

Nikolova, L. S. and Metzstein, M. M. (2015). Intracellular lumen formation in Drosophila proceeds via a novel subcellular compartment. *Development* **142**, 3964–3973.

Nüsslein-Volhard, C., Wieschaus, E. and Kluding, H. (1984). Mutations affecting the pattern of the larval cuticle in Drosophila melanogaster. I. Zygotic loci on the second chromosome. *Roux's Arch. Dev. Biol.* **193***, 267-282.*

Quesada, V., Conde, L., Villamor, N., Ordóñez, G. R., Jares, P., Bassaganyas, L., Ramsay, A. J., Beà, S., Pinyol, M., Martínez-Trillos, A., et al. (2011). Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nature genetics* **44**, 47–52.

Samakovlis, C., Hacohen, N., Manning, G., Sutherland D.C., Guillemin K., Krasnow M.A. (1996). Development of the Drosophila tracheal system occurs by a series of morphologically distinct but genetically coupled branching events. *Development* **122***, 1395-1407.*

Shaner, N. C., Lambert, G. G., Chammas, A., Ni, Y., Cranfill, P. J., Baird, M. A., Sell, B. R., Allen, J. R., Day, R. N., Israelsson, M., et al. (2013). A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum. *Nat. Methods* **10**, 407–409.

Shiga, Y., Tanaka-Matakatsu, M. and Hayashi, S. (1996). A nuclear GFP/beta-galactosidase fusion protein as a marker for morphogenesis in living Drosophila. *Development, Growth and Differentiation* **38**, 99–106.

Sutherland, D., Samakovlis, C. and Krasnow, M. A. (1996). branchless encodes a Drosophila FGF homolog that controls tracheal cell migration and the pattern of branching. *Cell* **87**, 1091–1101.

Urano, Y., Iiduka, M., Sugiyama, A., Akiyama, H., Uzawa, K., Matsumoto, G., Kawasaki, Y. and Tashiro, F. (2006). Involvement of the mouse Prp19 gene in neuronal/astroglial cell fate decisions. *J. Biol. Chem.* **281**, 7498–7514.

van der Lelij, P., Stocsits, R. R., Ladurner, R., Petzold, G., Kreidl, E., Koch, B., Schmitz, J., Neumann, B., Ellenberg, J. and Peters, J. M. (2014). SNW1 enables sister chromatid cohesion by mediating the splicing of sororin and APC2 pre-mRNAs. *EMBO J.* **33**, 2643–2658.

Wilusz, C. J., Wang, W. and Peltz, S. W. (2001). Curbing the nonsense: the activation and regulation of mRNA surveillance. *Genes & Development* **15**, 2781–2785.

Zhang, C., Li, W.-H., Krainer, A. R. and Zhang, M. Q. (2008). RNA landscape of evolution for optimal exon and intron discrimination. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 5797–5802.

Figures

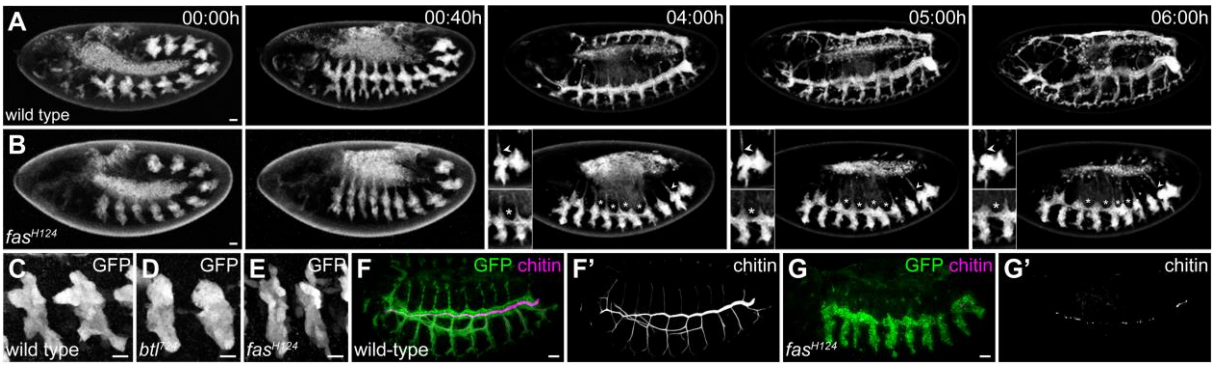


Figure 1

***H124* mutants show impaired primary tracheal branch outgrowth.**

(**A,B**) Time-lapse imaging of wild-type (**A**) and *fas*^{H124} (**B**) embryos (n=5). Tracheal cells are labeled with palmitoylated mNeonGreen. Primary branch outgrowth is reduced in *fas* embryos, although most metameres extend dorsal branches, which occasionally break off (arrowheads). Partial dorsal trunk connections are indicated by asterisks. (**C-E**) Stage 12 wild-type (**C**), *btl* (**D**) and *fas*^{H124} (**E**) embryos expressing cytosolic GFP in tracheal cells (n=5). Note that unlike in *btl* embryos (**D**) branching in *fas* embryos (**E**) is not completely abolished. (**F-G'**) Stage 14 wild-type (**F,F'**) and *fas*^{H124} (**G,G'**) embryos expressing GFP in tracheal cells and stained for chitin (n=8). Chitin is deposited in partial dorsal trunk connections in *fas* mutants.

Scale bars: 20 µm (**A,B,F-G'**); 10 µm (**C-E**).

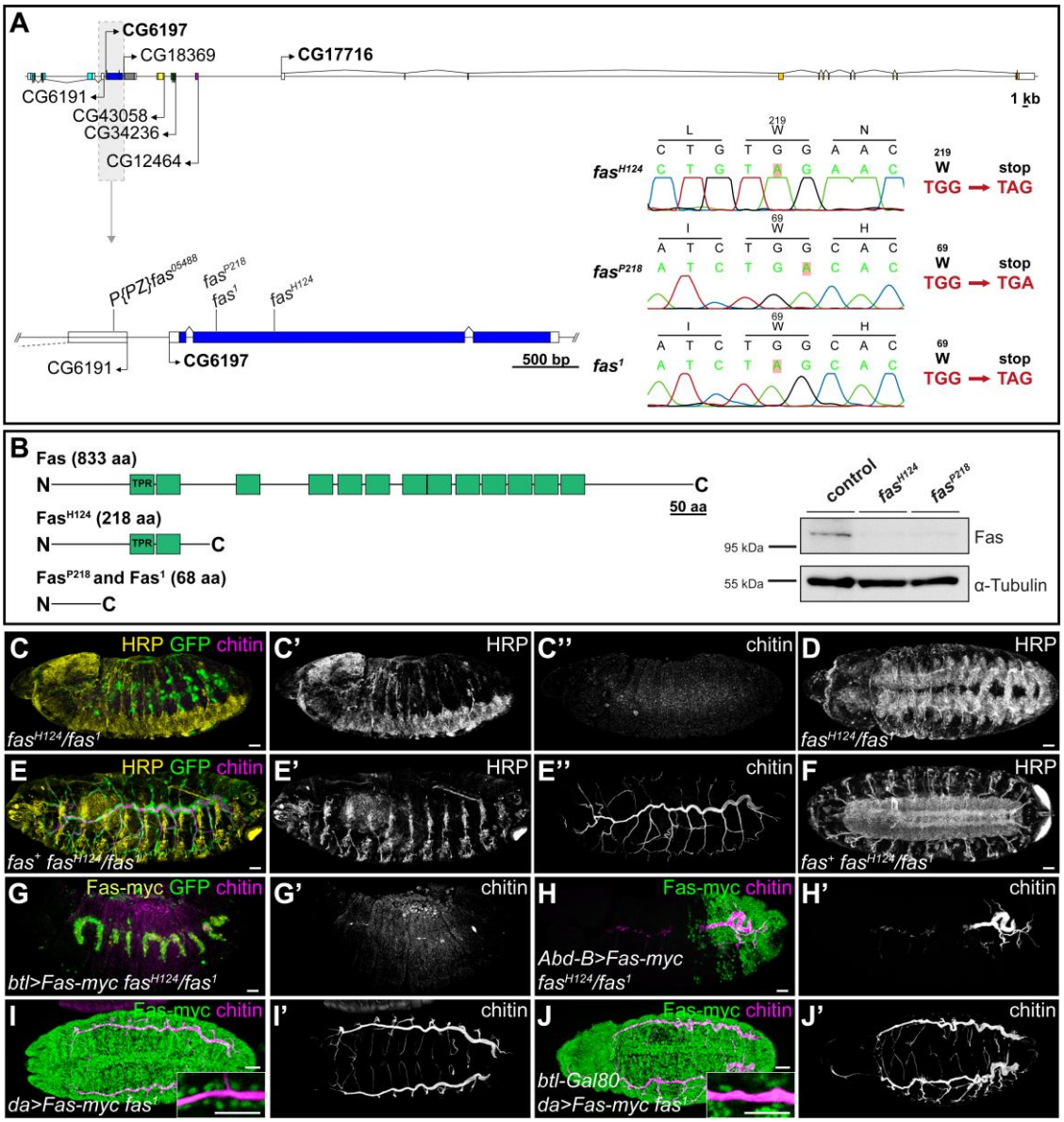


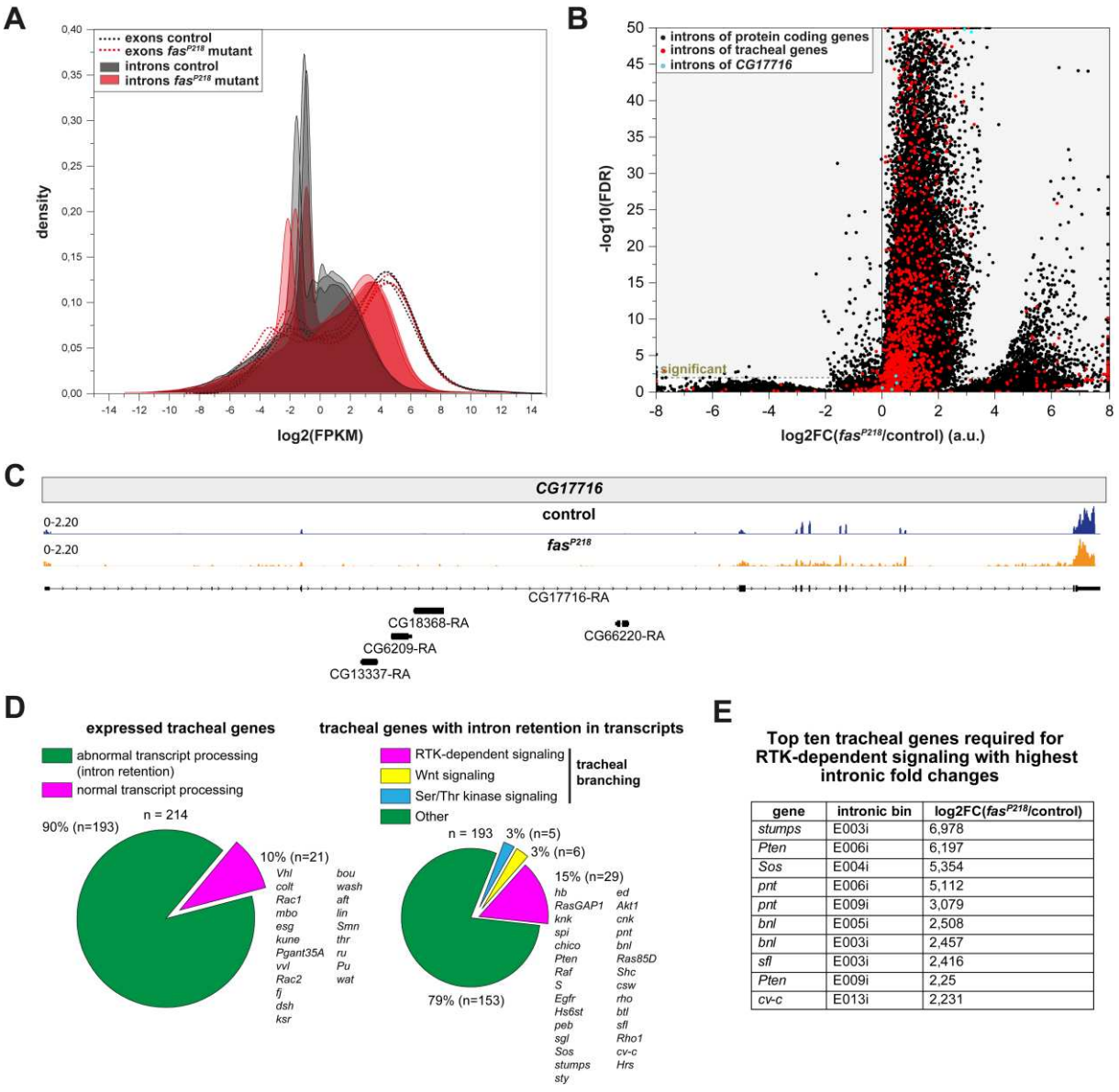
Figure 2

The NTC/Prp19C subunit *faint sausage* is required for tracheal branching.

(**A**) Structure of the *CG6179* (*fas*) locus. fas^{H124}, fas^{P218} and fas^{1} mutants carry premature stop-codons in *CG6179*. (**B**) Domain structure of Fas protein. *fas* mutants produce predicted truncated proteins lacking 11 (fas^{H124}) or all (fas^{P218} and fas^{1}) of the 13 tetratricopeptide repeats (TPR; predicted by TPRpred; https://toolkit.tuebingen.mpg.de/). Immunoblot from *fas* mutant and wild-type control embryos (15-18 hAEL) shows severely reduced Fas protein levels in *fas* mutants. (**C-F**) A genomic *fas* (*CG6179*) transgene (fas^{+}) rescues tracheal and

CNS defects of fas^{H124}/fas^{1} embryos (n=6). Lateral (**C-C'', E-E''**) and ventral (**D, F**) views of *fas* embryos (stage 15) lacking (**C-C'', D**) or carrying (**E-E'', F**) the fas^{+} construct. Embryos are stained for chitin (magenta) and HRP (yellow). Note that tracheal and CNS development are completely restored by the fas^{+} construct. (**G-H'**) fas^{H124}/fas^{1} embryos (stage 15) expressing cytosolic GFP (green) in tracheal cells and stained for chitin (magenta). Fas-myc (green) fails to rescue tracheal branching when expressed in tracheal cells using *btl*-Gal4 (n=8; **G, G'**), but rescues tracheal branching when expressed in the entire posterior body using *Abd-B*-Gal4 (n=7; **H, H'**). (**I-J'**) Ubiquitous expression of Fas-myc driven by *da*-Gal4 (**I-I'**) completely rescues tracheal development in fas^{1} embryos (stage 15). Tracheal development is largely rescued also when Fas-myc expression is blocked in tracheal cells by *btl*-Gal80 (n=3; **J-J'**). Insets show Fas-myc in nuclei of the tracheal DT (single plane).

Scale bars: 20 µm (**C-H'**), 30 µm (**I-J'**), insets 20 µm.



gene	intronic bin	log2FC(*fas*^{P218}/control)
stumps	E003i	6,978
Pten	E006i	6,197
Sos	E004i	5,354
pnt	E006i	5,112
pnt	E009i	3,079
bnl	E005i	2,508
bnl	E003i	2,457
sfl	E003i	2,416
Pten	E009i	2,25
cv-c	E013i	2,231

Figure 3

Zygotic *fas* function is required for efficient mRNA splicing.

(**A**) Density plots showing abundance of intronic and exonic sequences across all transcripts in *fas*^{P218} and control embryos. Intronic sequences accumulate in *fas*^{P218} mutants. (**B**) Intron retention. A large number of introns reveal significant (FDR < 0.01) retention in *fas*^{P218} embryos. The separate (right-hand) cluster of data points with large positive fold-changes mainly corresponds to intron bins with low intron inclusion levels in at least one condition. (**C**) Coverage plots of *CG17716* transcript. Note the presence of intronic reads in *fas*^{P218}, but

not in the control. (**D**) Pie charts showing abnormal transcript processing for tracheal genes in *fas* mutants. Left: Tracheal genes with abnormal transcript processing. 90% of the expressed genes with functions in tracheal development reveal significant intron retention (FDR < 0.01). Right: Functional classification of tracheal transcripts showing intron retention. Genes required for RTK signaling are listed. (**E**) Top ten tracheal genes involved in RTK signaling, ranked by intron retention fold-changes.

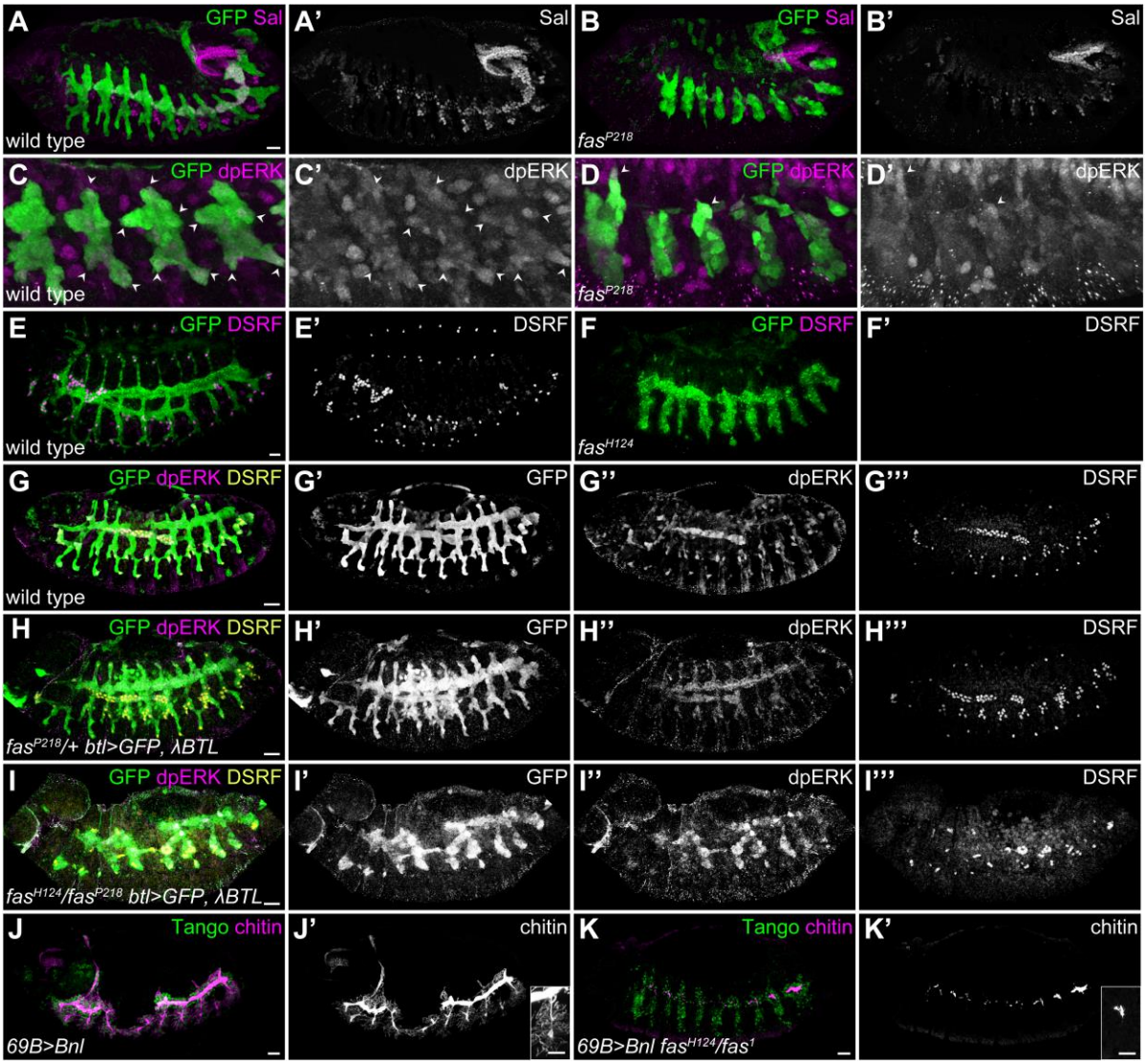


Figure 4

Fas is required for FGF and EGF signaling.

(**A-B'**) Stage 12 wild-type (**A,A'**) and *fas*^{P218} (**B,B'**) embryos expressing cytosolic GFP (green) in tracheal cells and stained for Sal (magenta). *fas* embryos show severely reduced Sal expression in tracheal cells. (**C-D'**) Stage 12 wild-type (**C,C'**) and *fas*^{P218} (**D,D'**) embryos stained for dpERK (magenta). dpERK accumulates at tracheal branch tips (arrowheads). Note that *fas* embryos lack dpERK accumulation in tracheal cells. (**E-F'**) Stage 14 wild-type (**E,E'**) and *fas*^{H124} (**F,F'**) embryos stained for stained for GFP (green) and DSRF (magenta). *fas* embryos lack DSRF-positive terminal cells. (**G-I**) Stage 14 wild-type (**G,G'**), *fas*

heterozygous (**H**) and *fas* homozygous (**I**) embryos expressing GFP (**G-I**) and constitutively active FGFR (λBtl; **H,I**) in tracheal cells, stained for GFP (green), dpERK (magenta) and DSRF (yellow). Note that tracheal dpERK accumulation (compare **D'** to **I''**) and DSRF expression (compare **F'** to **I'''**) are partially restored upon λBtl expression. (**J-K**) Stage 15 control (**J,J'**) and fas^{H124}/fas^{1} embryos (**K,K'**) misexpressing Bnl in the epidermis under the control of 69B-Gal4. Note that Bnl misexpression causes excessive tracheal branching in control, but not in *fas* embryos (compare **J'** to **K'**).

Scale bars: 20 µm (**A-B', E-F', G-K'**); 10 µm (**C-D'**); 15 µm insets (**J', K'**).

Supplementary Material

Supplementary Materials and Methods

Isolation and mapping of *fas* mutants

Mutations were induced by feeding EMS (25 mM) to males carrying a second chromosome marked with *btl*-Gal4, UAS-GFP and UAS-Verm-RFP transgenes. Mutagenized chromosomes were balanced over the *CyO Dfd-GMR-nvYFP* chromosome (Le *et al*., 2006). Living F3 embryos were analyzed for tracheal morphology and the distribution of Verm-RFP protein (Förster *et al*., 2010). Mutations were classified according to their phenotypes. Within each phenotypic class, mutants were crossed to each other to test for complementation of lethality.

The lethality of *H124* and *P218* mutations was mapped to the cytological interval 49C1-50D2 by non-complementation of the chromosomal deficiency *Df(2R)CX1*. Within this interval, 14 overlapping deficiencies (FlyBase) complemented *H124* and *P218* mutants. A small interval (50B4-B6) comprising six annotated genes was not covered by these deficiencies and contained a lethal P-element insertion (*P{PZ}05488*), which was allelic to the *fas* locus (Lekven *et al*., 1998) and failed to complement *H124* and *P218* mutants.

Sequencing of *fas* alleles

Genomic DNA was extracted from homozygous fas^{H124}, fas^{P218} and fas^{1} embryos selected by the absence of the *CyO Dfd-YFP* balancer chromosome (Le *et al*., 2006). Coding sequences of *CG6197* and *CG17716* were amplified and sequenced using oligonucleotides listed in Supplementary Table 3.

Statistics and reproducibility

For phenotypic analyses, sample size (*n*) was not defined using statistical methods, but was determined by taking into account the variability of a given phenotype. Investigators were not

blinded to allocation during experiments and samples were not randomised for the allocation of samples to experimental and control groups.

RNASeq data analysis

For differential gene expression, kallisto (v 0.42.1; Bray *et al*., 2016) was used to estimate the read count for each annotated transcript in each sample, and the counts were then summarized on the gene level. Genes with an estimated CPM (counts per million) exceeding one for at least three samples were retained for differential expression analysis and were considered as 'expressed'. The GLM framework of the edgeR R package (v 3.10.2; McCarthy *et al*., 2012; Robinson *et al*., 2010) was applied to test for differential gene expression between the control and mutant groups, using the sample preparation batch as a confounding factor in the model.

DEXSeq (Anders *et al*., 2012) was used to investigate the extent of intron retention. The reads were aligned to the *Drosophila melanogaster* genome (Ensembl v70) with STAR (v2.4.2a) (Dobin *et al*., 2013), retaining only reads aligning to a unique location. Based on the kallisto results obtained as described above, transcripts contributing less than 5% to the total abundance (TPM) of the corresponding gene in all samples were removed from the gtf file (Soneson *et al*., 2016). The reduced gtf file was processed with the python scripts provided with DEXSeq to "flatten" the annotation file by generating disjoint exon bins. Finally, the flattened annotation file was extended with intronic bins, defined as intragenic regions that did not overlap with any exon of the retained transcripts. Using this extended annotation, DEXSeq was used to quantify the abundance of each exonic and intronic bin, and to test each bin for differential inclusion between the control and mutant groups. Only results for intronic bins were retained for interpretation.

FPKMs were estimated for each exonic and intronic bin by dividing the normalized counts obtained from DEXSeq (adding 1 to avoid taking the log of 0) by the width of the bin and the total number of counts (exonic and intronic) for the corresponding sample.

The RNA-seq data has been deposited in ArrayExpress under accession number E-MTAB-5069.

Supplementary References

Anders, S., Reyes, A. and Huber, W. (2012). Detecting differential usage of exons from RNA-seq data. *Genome Res.* **22**, 2008–2017.

Bray, N. L., Pimentel, H., Melsted, P. and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21.

Förster, D., Armbruster, K. and Luschnig, S. (2010). Sec24-dependent secretion drives cell-autonomous expansion of tracheal tubes in Drosophila. *Curr. Biol.* **20**, 62–68.

Le, T., Liang, Z., Patel, H., Yu, M. H., Sivasubramaniam, G., Slovitt, M., Tanentzapf, G., Mohanty, N., Paul, S. M., Wu, V. M., et al. (2006). A new family of Drosophila balancer chromosomes with a w-dfd-GMR yellow fluorescent protein marker. *Genetics* **174**, 2255–2257.

Lekven, A. C., Tepass, U., Keshmeshian, M. and Hartenstein, V. (1998). faint sausage encodes a novel extracellular protein of the immunoglobulin superfamily required for cell migration and the establishment of normal axonal pathways in the Drosophila nervous system. *Development* **125**, 2747–2758.

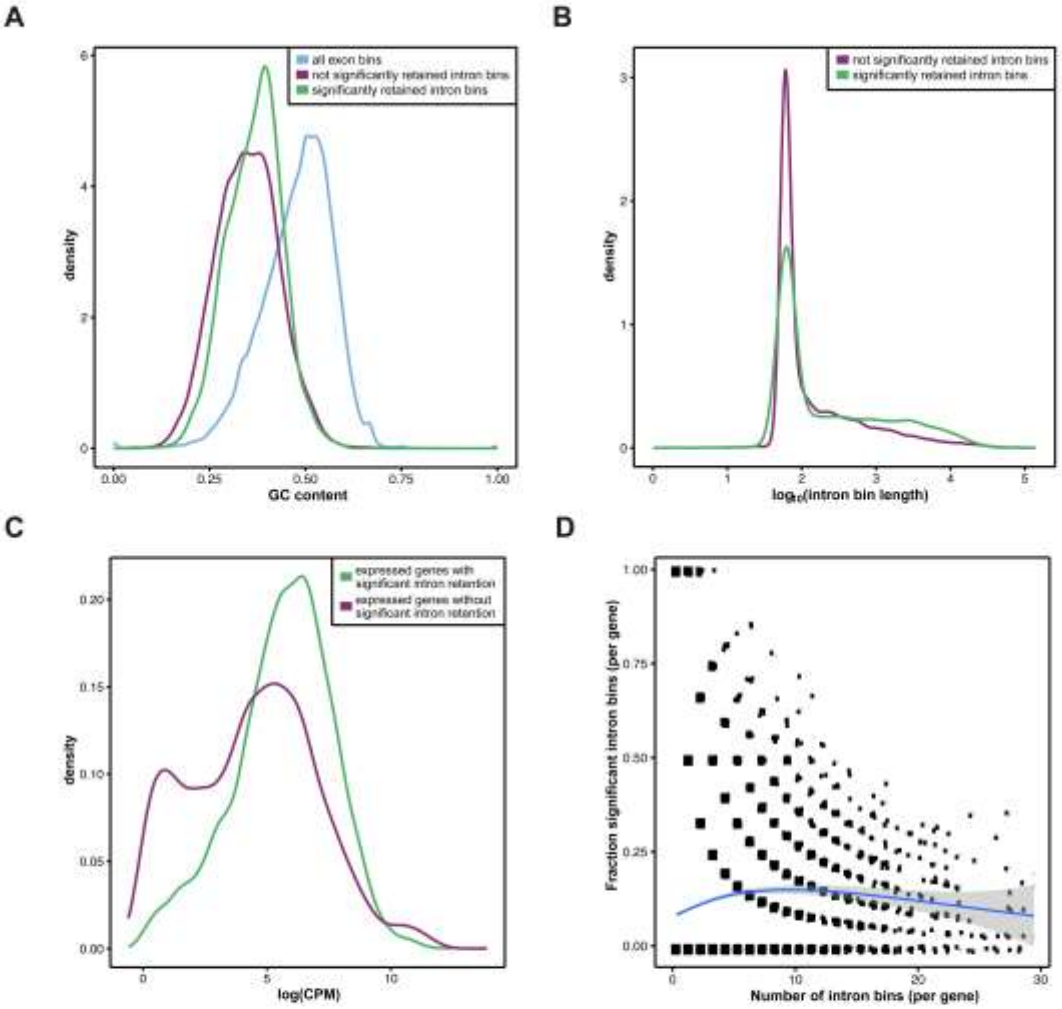
McCarthy, D. J., Chen, Y. and Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research* **40**, 4288–4297.

Robinson, M. D., McCarthy, D. J. and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140.

Soneson, C., Matthes, K. L., Nowicka, M., Law, C. W. and Robinson, M. D. (2016). Isoform prefiltering improves performance of count-based methods for analysis of differential transcript usage. *Genome Biol.* **17**, 248.

Supplementary Figures

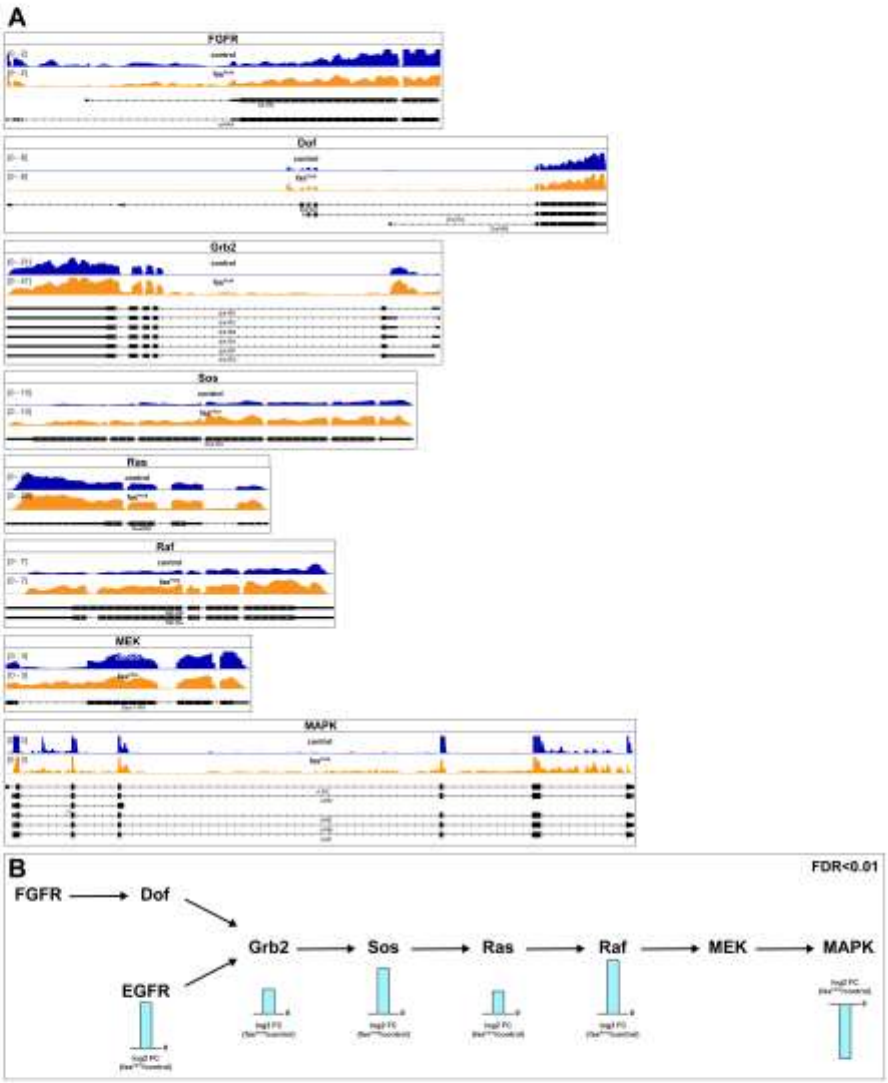
Sauerwald_Supplementary_figure_1



Supplementary Figure 1

Characterization of significantly retained intron bins and the genes containing them. (**A**) Distribution of GC content for the significantly retained intron bins (adjusted $p < 0.01$, log2FC > 1.5), not significantly retained intron bins (adjusted $p > 0.1$) and all exon bins. Significantly retained intron bins show slightly increased GC content compared to not significantly retained intron bins. (**B**) Distribution of the bin lengths of significantly retained intron bins (adjusted $p < 0.01$, log2FC > 1.5) and not significantly retained intron bins (adjusted $p > 0.1$). The significantly retained intron bins are slightly enriched with long bins. This association may potentially be explained by the higher power of count-based methods to give significant test results for features with large read counts, since longer bins give rise to more fragments than short bins, and thus a higher number of reads. (**C**) Distribution of expression values for expressed genes (i.e., genes not filtered out in the differential gene expression analysis) containing any of the significantly retained introns or not. Similarly to (B), the difference may be partly explained by the higher power of count-based methods to detect differences in features with high numbers of counts. (**D**) The association between the number of intron bins per gene and the fraction of these showing significant retention (adjusted $p < 0.01$, log2FC > 1.5). Each dot represents a gene, and dots have been jittered slightly to decrease overplotting. The blue line represents a smooth fit to the points. No strong association between the expected fraction of significant bins and the number of bins is apparent.

Sauerwald_Supplementary_Figure_2



Supplementary Figure 2

Abnormal transcript processing of receptor tyrosine kinase signaling components.

(**A**) Coverage plots of transcripts encoding components of EGF and FGF signaling. (**B**) Significant (FDR < 0.01) changes in transcript levels of EGF and FGF signaling components.

Supplementary Table 1

Analysis of intron retention on intron-bin-level using DEXSeq. Columns include the gene identifiers, gene names, details about the corresponding intronic bins, fold changes (FC) for intron retention and statistics relating to DEXSeq analysis.

Click here to Download Table S1

Supplementary Table 2

Gene-level differential expression analysis using kallisto and edgeR. Columns include the gene identifiers, gene names, fold changes (FC) and statistics relating to differential expression analysis.

Click here to Download Table S2

Supplementary Table 3

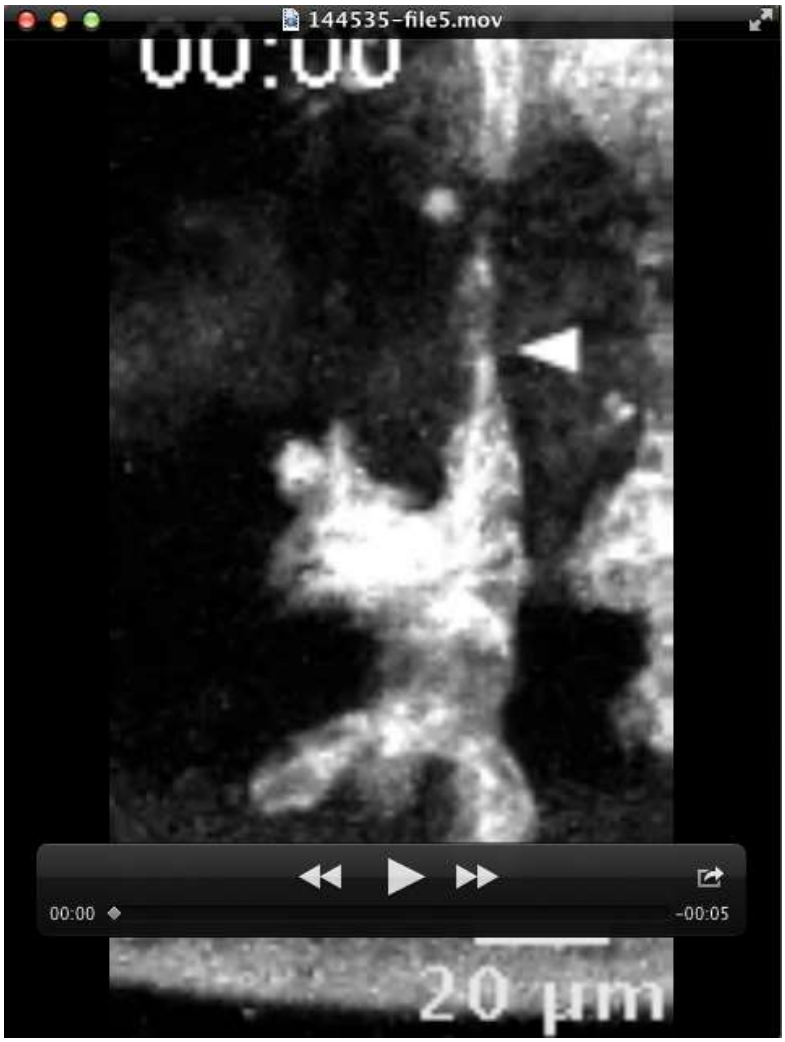
List of oligonucleotides used for sequencing of *CG6197* and *CG17716* coding region.

primer	sequence	description
JuS15	ACTTTGGTATTAGCTGCGGTA	sequencing *CG6197* exon 1-2
JuS16	CGCACAATTCGTTCCACAGC	sequencing *CG6197* exon 1-2
JuS17	CTGCGAGTCTACCGTCGATA	sequencing *CG6197* exon 2
JuS18	GCACCGTTTCCGTATCATCA	sequencing *CG6197* exon 2
JuS19	GGAGTTCGCCAAGTTCTACG	sequencing *CG6197* exon 2
JuS20	CACTTCGCCCAACTTCGTTT	sequencing *CG6197* exon 2
JuS21	AAGGCAGCCGAGATCTATGG	sequencing *CG6197* exon 2-3
JuS22	TCTTGGATTGCGCGTAAAAGC	sequencing *CG6197* exon 2-3
SL79	GTCTTTGTTTGCCCTGCACT	sequencing *CG17716* exon 3
SL80	ACCTGCCTGACTGCGATTAT	sequencing *CG17716* exon 3
JuS9	GGGACCAAGGAATTGCAGTG	sequencing *CG17716* exon 4
JuS10	TGAGACAGACGGAGGGAAAA	sequencing *CG17716* exon 4
JuS7	TTGCCACATTACCAGTGGAT	sequencing *CG17716* exons 5-6
JuS8	GCTCGAGTGCAAGACAAACA	sequencing *CG17716* exons 5-6
SL71	TACACCCCCGATGATTGATT	sequencing *CG17716* exon 7
SL72	GATACCCAAGCGCACTGTTT	sequencing *CG17716* exon 7
SL73	CACTCTGCGAGCCTAGCATT	sequencing *CG17716* exon 8-9
SL74	ATGTGCGGGTGTCGTAAAAT	sequencing *CG17716* exon 8-9
SL75	CAACAGCCCTCGGATTTG	sequencing *CG17716* exon 10-11
SL76	CTGATTCATCGATTGGGTTG	sequencing *CG17716* exon 10-11
SL77	ATGCGAGGCGAAACAAACT	sequencing *CG17716* exon 12-13
SL78	GACTAGGGCAATTTGGATGC	sequencing *CG17716* exon 12-13

Supplementary Movie 1

Tracheal branch outgrowth is affected in *fas*^{H124} mutant embryo.

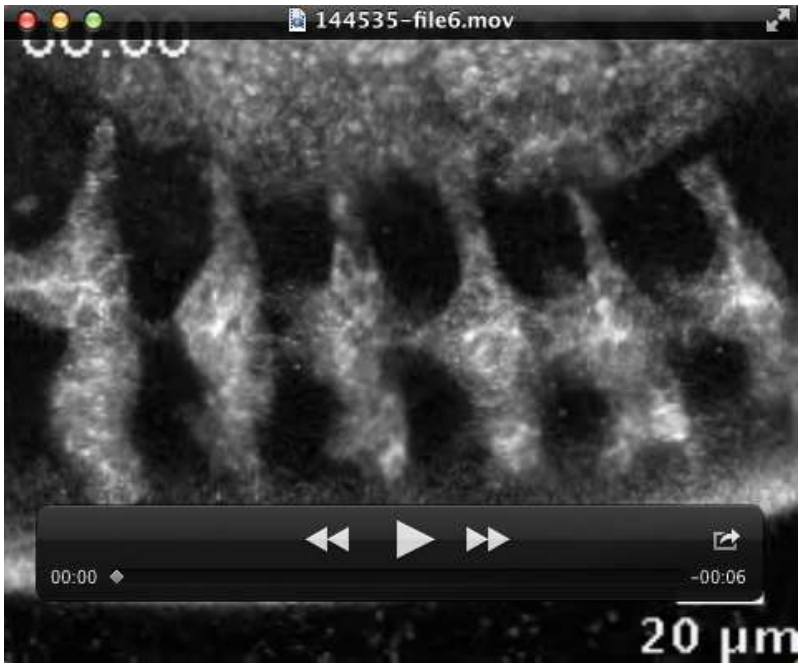
Time-lapse movies of wild-type (top) and *fas*^{H124} homozygous mutant (bottom) embryos expressing palmitoylated mNeonGreen under the control of *btl*-Gal4. The movies were acquired with a 20x/0.75 NA objective and a frame rate of 5 min.



Supplementary Movie 2

Dorsal branch separating from the tracheal primordium in *fas*^{H124} mutant embryo.

Time-lapse movie of tracheal metamere one in a *fas*^{H124} homozygous mutant embryo expressing palmitoylated NeonGreen under the control of *btl*-Gal4. The movie was acquired with a 20x/0.75 NA objective and a frame rate of 5 min.



Supplementary Movie 3

Formation of dorsal trunk connections in *fas*^{H124} mutant embryo.

Time-lapse movie of tracheal metameres forming DT connections in a *fas*^{H124} homozygous mutant embryo expressing palmitoylated NeonGreen under the control of *btl*-Gal4. The movie was acquired with a 20x/0.75 NA objective and a frame rate of 5 min.